КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ
COMPUTER SIMULATION

УДК 004.8:343.721
DOI 10.52575/2687-0932-2026-53-1-111-121
EDN ITSPKU

Применение методов машинного обучения для выявления мошенничества в банковских транзакциях

Хамитов Р.М., Куценко С.М., Салтанаева Е.А.
Казанский государственный энергетический университет,
Россия, 420066, Республика Татарстан, г. Казань, ул. Красносельская, д. 51
hamitov@gmail.com, s.koutsenko@mail.ru, elena_maister@mail.ru

Аннотация. Статья затрагивает связь массового распространения интеллектуальных технологий и увеличения количества возможных противоправных действий с их использованием. Рассматривается разработка и анализ методов машинного обучения для выявления мошеннических транзакций в финансовой сфере. Актуальность темы обусловлена постоянной эволюцией мошеннических схем, что требует применения инновационных технологий для обеспечения финансовой безопасности. В работе рассматриваются современные подходы к обработке данных, масштабирование и устранение асимметрии данных. Исследование охватывает обучение моделей с использованием четырех алгоритмов: логистической регрессии, дерева решений, метода случайного леса и градиентного спуска. Для оценки качества модели была использована метрика ROC-AUC, а также такие характеристики, как точность, полнота и F1-мера. Лучшие результаты показала модель логистической регрессии, достигнув значения ROC-AUC 0,975 на тестовом наборе данных. Результаты работы подчеркивают практическую ценность моделей машинного обучения как надежного инструмента для минимизации рисков, связанных с мошенническими транзакциями.

Ключевые слова: мошеннические транзакции, интеллектуальные системы, методы машинного обучения, точность модели, эффективность алгоритма, кросс-валидация

Для цитирования: Хамитов Р.М., Куценко С.М., Салтанаева Е.А. 2026. Применение методов машинного обучения для выявления мошенничества в банковских транзакциях. *Экономика. Информатика*, 53(1): 111–121. DOI 10.52575/2687-0932-2026-53-1-111-121. EDN ITSPKU

Application of Machine Learning Methods to Detect Fraud in Bank Transactions

Renat M. Khamitov, Svetlana M. Kutsenko, Elena A. Saltanaeva
Kazan State Power Engineering University,
51 Krasnoselskaya St., Kazan 420066, Tatarstan, Russia
hamitov@gmail.com, s.koutsenko@mail.ru, elena_maister@mail.ru

Abstract. The article touches on the connection between the mass distribution of intellectual technologies and their possible increased use in illegal actions. The authors focus on development and analysis of machine learning methods for detecting fraudulent transactions in the financial sector. The relevance of the topic is explained by the ongoing evolution of fraudulent schemes, which requires innovative technologies to be

applied for ensuring financial security. The paper considers modern approaches to data processing, scaling and elimination of data asymmetry. The study covers model training using four algorithms: logistic regression, decision tree, random forest method, and gradient descent. To assess the quality of the model, the ROC-AUC metric was used, as well as characteristics such as accuracy, completeness and F1 measure. The logistic regression model performed best, achieving a ROC-AUC value of 0.975 on the test dataset. The results of the work highlight the practical value of machine learning models as a reliable tool for minimizing the risks associated with fraudulent transactions.

Keywords: fraudulent transactions, intelligent systems, machine learning methods, model accuracy, algorithm efficiency, cross-validation

For citation: Khamitov R.M., Kutsenko S.M. Saltanaeva E.A. 2026. Application of Machine Learning Methods to Detect Fraud in Bank Transactions. *Economics. Information technologies*, 53(1): 111–121 (in Russian). DOI 10.52575/2687-0932-2026-53-1-111-121. EDN ITSPKU

Введение

В настоящее время использование интеллектуальных технологий и систем приобретает статус неотъемлемой части во всех сферах как деятельности государства, так и жизни каждого отдельного гражданина. К сожалению, при массовом внедрении и использовании новых интеллектуальных технологий с их помощью злоумышленниками совершаются также и противоправные действия, требующие противодействия. Наиболее часто подвергается атакам финансовая сфера.

Мошенничество в банковском секторе имеет давнюю историю, начиная с формирования первых банковских учреждений. Ситуацию усугубляет возможность использования развитых интеллектуальных технологий также и злоумышленниками для реализации мошеннических схем. Мошенничество, связанное с транзакциями в финансовой сфере, подрывает доверие к финансовой системе государства в целом. В виду того, что схемы мошенничества модернизируются, усложняются и совершенствуются постоянно, требуется непрерывное улучшение методов борьбы с ними для обеспечения финансовой безопасности. На сегодняшний день существуют различные технологии, способные определять мошеннические транзакции и предотвращать их проведение [Omolara et al., 2024; Ioffe, 2024; Хлобыстова, Абрамов, 2024]. К таким технологиям можно отнести: анализ больших данных; биометрические технологии; многофакторную аутентификацию; системы оценки риска в реальном времени; машинное обучение. Технологию Big Data для определения мошеннических операций использует малая часть банков России, в частности, Т-Банк и Сбербанк, поскольку технология требует значительных ресурсов и вычислительных мощностей, а также внимательного подхода к защите данных и управлению системами. Использование биометрии в рамках безопасности для определения мошеннической транзакции продемонстрировал Почта Банк. В банковской сфере многофакторная аутентификация, как и биометрическая технология, используется в совокупности с другими методами определения мошеннических транзакций. Определение мошеннических транзакций с помощью методов машинного обучения становится главным вызовом в защите интересов не только клиентов, но и компаний [Zhu, Zhou, 2023; Ye, 2023; Chio, 2020; Wang et al., 2023]. Платформа FICO Falcon Fraud Manager является мощным инструментом для решения задач организации по обнаружению мошенничества при транзакциях [ICO Falcon Fraud Manager]. Однако многие крупные банковские и финансовые учреждения стремятся к созданию собственной системы определения мошенничества, используя передовые технологии, в частности методы машинного обучения [Мартин, 2022]. Определение мошеннических транзакций методами машинного обучения имеет ряд преимуществ по сравнению с традиционными методами, основанными на правилах: точность; скорость; эффективность [Аскаров, Хамитов, 2024]. Точность методов машинного обучения намного выше в виду того, что они не ограничены определенными правилами и условиями, наоборот, алгоритмы

пытаются найти сходства в случаях мошенничества и выстроить общую закономерность [Траск, 2022]. Время определения мошенничества методами машинного обучения сокращается за счет увеличения количества транзакций, поскольку модели машинного обучения самообучаются эффективнее, позволяя находить более нестандартные и менее связанные случаи мошенничества.

Целью работы было подготовить, обучить и проанализировать модель определения мошеннических транзакций, способную эффективно выявлять аномалии и предотвращать мошенничество в банковских транзакциях.

Объект и методы исследования

Инструментами разработки были выбраны язык программирования Python и среда разработки Google Colab, способные обеспечить эффективность, масштабируемость и гибкость. Интеграция Google Colab и Python позволяет эффективно решать задачи, связанные с машинным обучением, обрабатывать большие массивы данных и легко масштабировать проекты, что является ключевыми требованиями для разрабатываемой модели [Марченко, 2023; Madani, 2023; Yuxi, 2020]. Облачная среда Google Colab основана на платформе Jupyter Notebook Framework, что делает ее удобным инструментом для работы с Data science и машинным обучением. Выбранными библиотеками в Colab стали предустановленные библиотеки, такие как Tensorflow, PyTorch, Keras, Sklearn [Орельен, 2020]. Colab подходит для машинного обучения и анализа данных, предоставляя пользователям бесплатный доступ к мощным вычислительным ресурсам, таким как GPU (графический процессор) и TPU (тензорный процессор), необходимым для быстрого и эффективного обучения [Григорьев, 2023; Kelleher, 2019]. Помимо своей простоты, гибкости, обширной поддержки Python обладает огромным количеством библиотек для машинного обучения [Плас, 2021]. Среди наиболее известных и популярных библиотек для разработки и обучения модели выделяют: XGBoost; Scikit-learn (sklearn); Matplotlib; Seaborn; Pandas; NumPy; SciPy. Использование этих библиотек облегчило и ускорило работу на языке программирования Python, в результате чего получилось достичь поставленных целей в разработке модели машинного обучения для определения мошеннических транзакций.

В рамках обучения модели использовался набор данных, содержащий 284 807 помеченных транзакций, из которых 492 транзакции являлись мошенническими. Столь малое количество мошеннических транзакций говорит о том, что датасет имеет высокую степень несбалансированности – на позитивные классы (мошеннические) приходится 0,172 % от общего числа транзакций [Окуньков и др., 2023].

Перед составлением набора данных все транзакции были обезличены и преобразованы в числовые значения с помощью анализа главных компонентов (PCA). Основная цель данного анализа заключается в том, чтобы уменьшить размерность используемого набора данных, при этом сохранив наиболее важные закономерности между переменными. В методах классификации и регрессии переменные, которые представляют избыточную информацию, могут снизить точность модели, вследствие чего уменьшение размерности входных данных позволяет качественнее и эффективнее обучить модель. В результате преобразования датасет состоит из 31 признака – 28 числовых признаков V1, V2, V3, … V28 и признаки Time, Amount, Class, которые не были преобразованы методом PCA. Признак Time обозначает время между первой и текущей транзакциями. Признак Amount говорит о том, на какую сумму совершена транзакция. Данный признак полезен в ситуациях, когда необходимо отталкиваться от суммы транзакции при поиске мошенничества. Признак Class принимает два значение – 1 в случаях, когда транзакция является мошеннической, и 0, когда транзакция является обычной.

Была построена матрица корреляции. Коэффициент корреляции принимает значения от -1 до 1, где положительные корреляции обозначают, что при увеличении одной переменной другая также растет, отрицательная же корреляция, наоборот, говорит о том, что при увеличении одной переменной другая имеет тенденцию к уменьшению. Значения, близкие

к -1 и 1, являющиеся сильной корреляцией, предполагают более выраженную связь между значениями, тогда как значения, близкие к нулю, обозначают менее выраженную связь. Построенная матрица корреляции свидетельствует о слабой корреляции, поскольку в ней не было признаков, превышающих 0,1.

Для обучения и тестирования модели определения мошеннических транзакций набор данных был разделен на обучающий и тестовый наборы в соотношении 80:20. В ходе проведенных манипуляций количество положительных транзакций в обучающем наборе составило 394, количество положительных транзакций в тестовом наборе – 98.

Следующим этапом стало масштабирование данных. Масштабирование проводилось относительно признака Amount с помощью метода RobustScaler библиотеки sklearn. Данный метод масштабирует элементы, используя статистику, устойчивую к выбросам, и удаляет медиану, тем самым масштабируя данные в диапазоне от 1-го квартиля до 3-го, то есть от 25-го квантиля до 75-го. Поскольку набор данных сильно несбалансирован, масштабирование методом RobustScaler позволило устранить влияние выбросов по причине того, что выбросы могут исказить общий характер данных и негативно сказаться на производительности модели. Масштабирование выполнено на тестовой и обучающей выборках.

Далее была проверена асимметрия данных. Распределение называется ассиметричным, если хвост с левой или с правой части более выражен, иными словами среднее значение, мода и медиана не совпадают. Выявление и исправление асимметрии в несбалансированном наборе помогло улучшить производительность модели. Для начала была проверена асимметрия с помощью метода skew() – если асимметрия находится в пределах от -1 до 1, то для преобразования данных будет использоваться силовое преобразование. В таблице показаны признаки датасета и их значения асимметрии. Для устранения асимметрии был использован метод PowerTransformer библиотеки sklearn, с помощью которого распределение было приведено к гауссовому виду, а затем был применен метод Йео – Джонсона для преобразования в симметричное распределение.

Асимметрия датасета
Dataset asymmetry

Признаки	Значения	Признаки	Значения	Признаки	Значения	Признаки	Значения
V1	-3.306334	V8	-8.639485	V15	-0.308419	V22	-0.219171
V2	-4.779484	V9	0.541869	V16	-1.077909	V24	-0.549854
V3	-2.247962	V10	1.132688	V17	-3.733377	V25	-0.436292
V4	0.687574	V11	0.354102	V18	-0.254948	V26	0.574980
V5	-2.786851	V12	-2.286654	V19	0.106133	V27	-0.890209
V6	1.937381	V13	0.064819	V20	-1.960492	V28	9.978409
V7	3.152665	V14	-1.699112	V21	3.490183	Итого	18.193943

В ходе обучения модели были использованы четыре алгоритма машинного обучения. Использование нескольких алгоритмов позволило сравнить их эффективность и определить лучшую модель для выявления мошеннических транзакций. Были рассмотрены алгоритмы: логистическая регрессия (Logistic Regression); дерево решений (Decision Tree); метод случайного леса (Random Forest); градиентный спуск (XGBoost).

В качестве оценки производительности модели была использована кросс-валидация K-fold. Кросс-валидация позволяет достигнуть баланса между переобучением и недообучением модели. В основе метода лежит разделение набора данных на часть, используемую для обучения модели, и другую часть – для её проверки. Стратифицированная кросс-валидация равномерно распределяет классы в каждом наборе данных. Это особенно важно в наборах данных, когда один класс составляет значительную часть датасета, что может привести к переобучению модели.

Для оценки качества модели будем использовать метрику ROC-AUC [Liu et al., 2024]. Кривая ROC – это график оценки модели бинарной классификации при различных пороговых значениях, отображающий частоту истинных положительных результатов (TPR) и ложных положительных результатов (FPR), в свою очередь AUC – это площадь под кривой ROC. Показатель TPR измеряет долю истинных положительных результатов, которые модель правильно классифицирует, показатель FPR измеряет долю ложных положительных результатов, которые модель неправильно классифицирует. ROC-AUC представляет из себя численный показатель качества классификатора, варьируясь от 0 до 1, где 0 – это случайное предсказание, а 1 – идеальное разделение классов. В нашем случае данные сильно несбалансированны (лишь 0,17% транзакций из общего количества транзакций являются мошенническими), в результате чего точность не будет подходящей метрикой для оценки модели. Использование максимального значения ROC-AUC полезно для ситуаций поиска наилучшей модели, однако использование среднего значения является более надежным вариантом в рамках использования кросс-валидации. На рис. 1 показан график ROC-AUC.

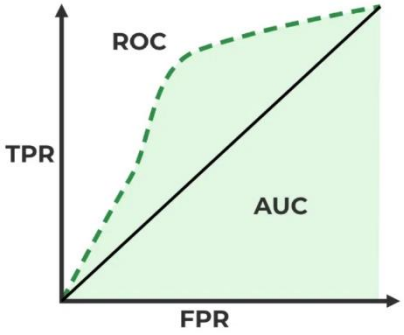

Рис. 1. График метрики ROC AUC
Fig. 1. ROC AUC metric plot

Результаты и их обсуждение

В рамках проекта было необходимо разработать концепцию системы распознавания лекарственных средств по фотографиям упаковок для дальнейшей интеграции в многофункциональный чат-бот.

В рамках разработки модели для обнаружения мошеннических транзакций в первую очередь необходимо выполнить стратифицированную кросс-валидацию K-Fold. В небольших наборах данных время обучения достаточное для того, чтобы проверить различные комбинации гиперпараметров. Однако в нашем случае необходимо использовать рандомизированный поиск ввиду случайной и неравномерной выборки.

При обучении модели логистической регрессии подбирается гиперпараметр C, который был использован для подбора параметров модели. Для каждого значения гиперпараметра производится кросс-валидация в следующем порядке: 1) разделение данных на обучающую и тестовую выборки; 2) обучение логистической регрессии; 3) получение прогнозов на основе обучающей выборки; 4) вычисление метрик – ROC-AUC, точность, полнота и F1-мера; 5) построение ROC-кривой.

По результатам обучения модели получили следующие значения:

– лучшее среднее значение ROC-AUC для гиперпараметра C: 0,979;

– лучшее среднее значение точности для гиперпараметра C: 0,885;

– лучшее среднее значение полноты для гиперпараметра C: 0,629;

– лучшее среднее значение F1-меры для гиперпараметра C: 0,734.

В результате настройки гиперпараметров с использованием K-кратной перекрестной проверки для логистической регрессии получили лучшее значение ROC-AUC, равное 0,981, и лучший гиперпараметр {“C”: 0,01, “penalty”: “l2”}.

На рис. 2 изображен график кривой ROC для модели логистической регрессии.

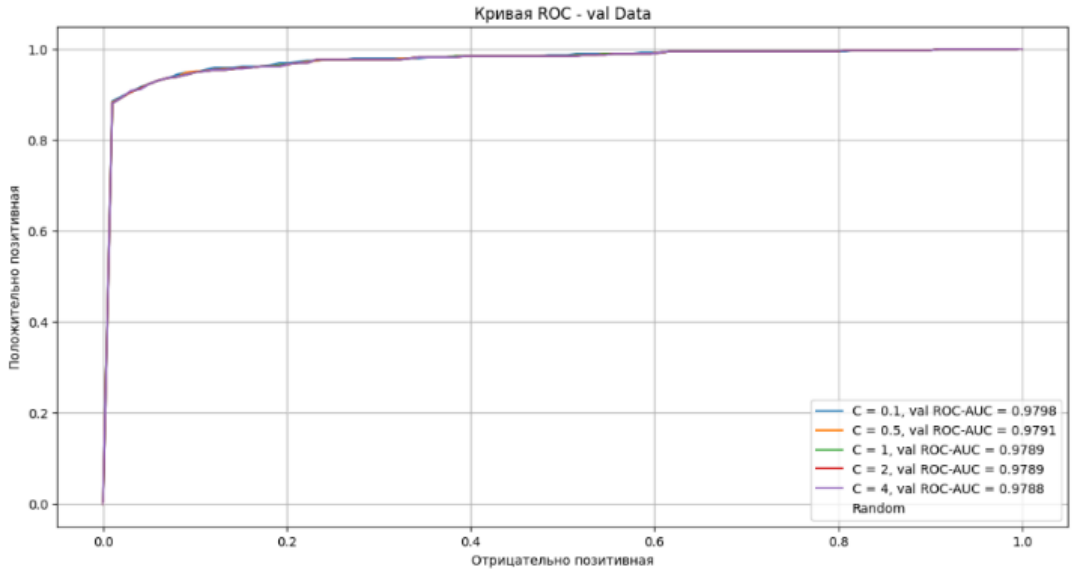

Рис. 2. Логистическая регрессия
Fig. 2. Logistic regression

При применении алгоритма Дерево решений (Decision Tree) получены следующие результаты обучения модели:

– лучшее значение Max Depth: 3;
– лучшее среднее значение ROC-AUC: 0,934;
– среднее значение точности лучшего значения max_depth: 0,848;
– среднее значение полноты лучшего значения max_depth: 0,716;
– среднее значение F1-меры лучшего значения max_depth: 0,775.

В результате настройки гиперпараметров получили лучшее значение ROC-AUC, равное 0.934, и лучший гиперпараметр {“criterion”: “entropy”, “max_depth”: 3, “min_samples_leaf”: 1, “min_samples_split”: 2}.

На рис. 3 изображен график кривой ROC для модели дерева решений.

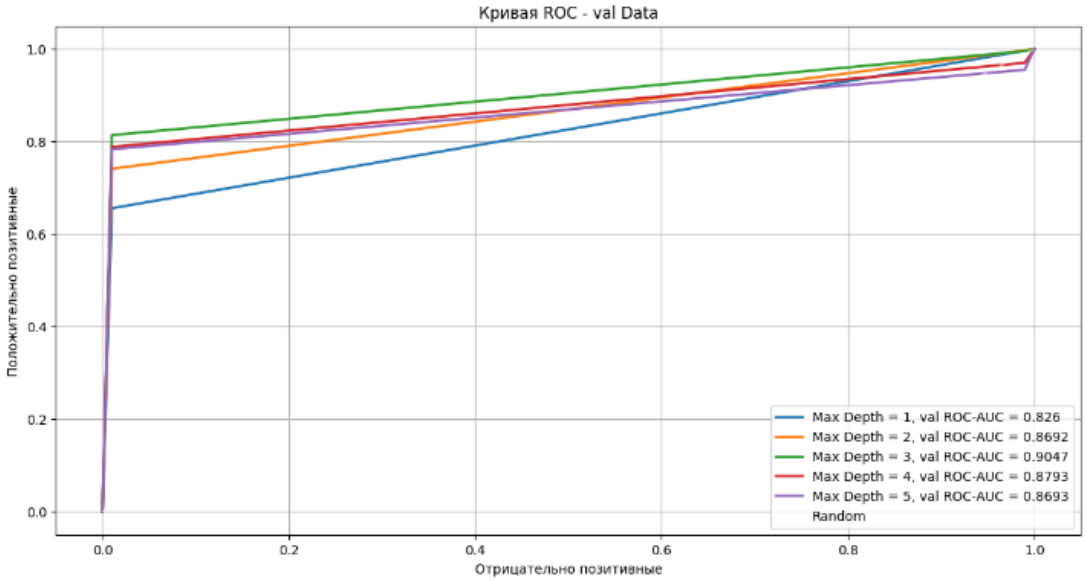

Рис. 3. Кривая ROC для дерева решений
Fig. 3. Decision Tree ROC Curve

При обучении модели методом случайного леса (Random Forest) получили следующие значения:

– средние значения ROC-AUC для train data при каждом значении n Estimators: [0,999, 1,0, 1,0, 1,0, 1,0];

– средние значения ROC-AUC для val data при каждом значении n Estimators: [0,927, 0,944, 0,946, 0,954, 0,959];

– лучшее значение n Estimators: 400;

– лучшее среднее значение ROC-AUC для val data: 0,959.

После выполнения настройки классификатора метода случайного леса получили лучшее значение ROC-AUC, равное 0,964, и лучшее значение гиперпараметров {“min_samples_split”: 7, “n_estimators”: 500}.

На рис. 4 показаны графики кривых ROC для метода случайного леса.

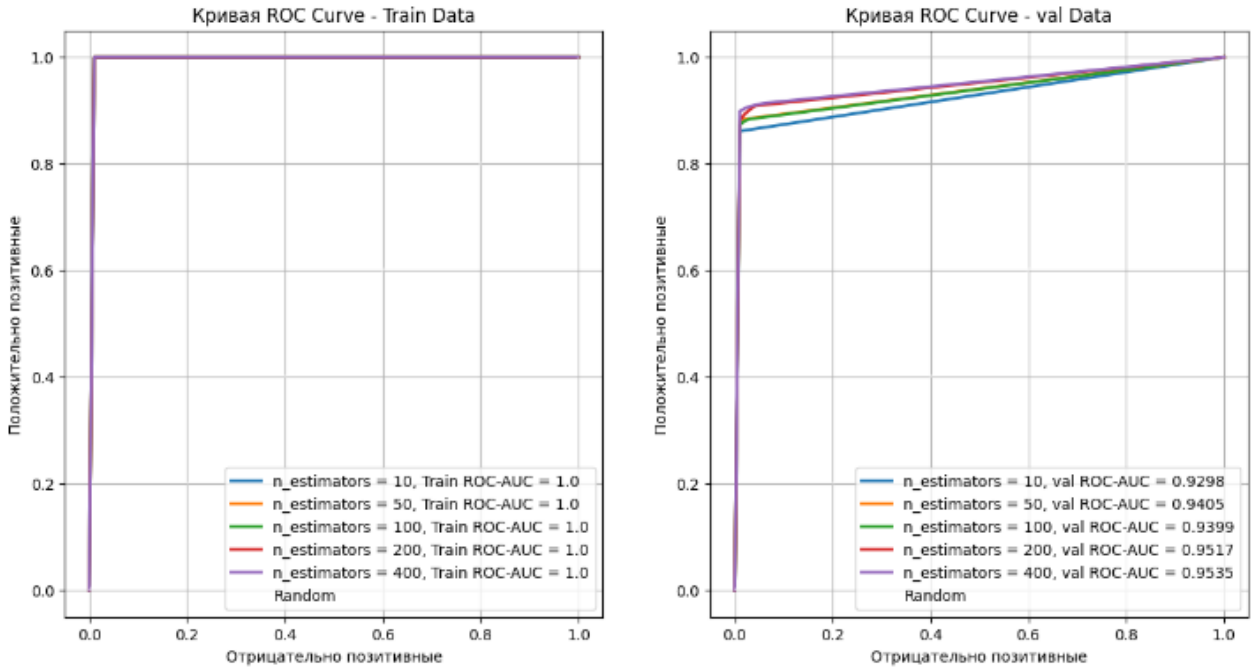

Рис. 4. Кривая ROC для метода случайного леса
Fig. 4. ROC curve for random forest method

Обучающий алгоритм XGBoost выдал следующие значения:

– средние значения ROC-AUC для val data при каждом значении Learning Rate: [0,929, 0,945, 0,984, 0,98, 0,966];

– лучшее значение Learning Rate: 0,1;

– лучшее среднее значение ROC-AUC для val data: 0,984;

– среднее значение точности val для лучшего значения learning_rate: 0,923;

– среднее значение отзыва val для лучшего значения learning_rate: 0,777;

– среднее значение F1 val для лучшего значения learning_rate: 0,843.

После выполнения настройки классификатора градиентного спуска получили лучшее значение ROC-AUC, равное 0,983, и лучшее значение гиперпараметров {“learning_rate”: 0,1, “max_depth”: 3, “subsample”: 0,7}.

На рис. 5 показаны графики кривых ROC для градиентного спуска.

Итак, в ходе обучения выбранными методами машинного обучения были получены следующие значения ROC-AUC и их лучшие гиперпараметры на основе используемого набора данных.

Перед началом проверки моделей на тестовом наборе необходимо масштабировать тестовый датасет по признаку Amount. Далее была проведена оценка эффективности примененных алгоритмов. Процесс эффективности методов машинного обучения рассмотрен на примере логистической регрессии.

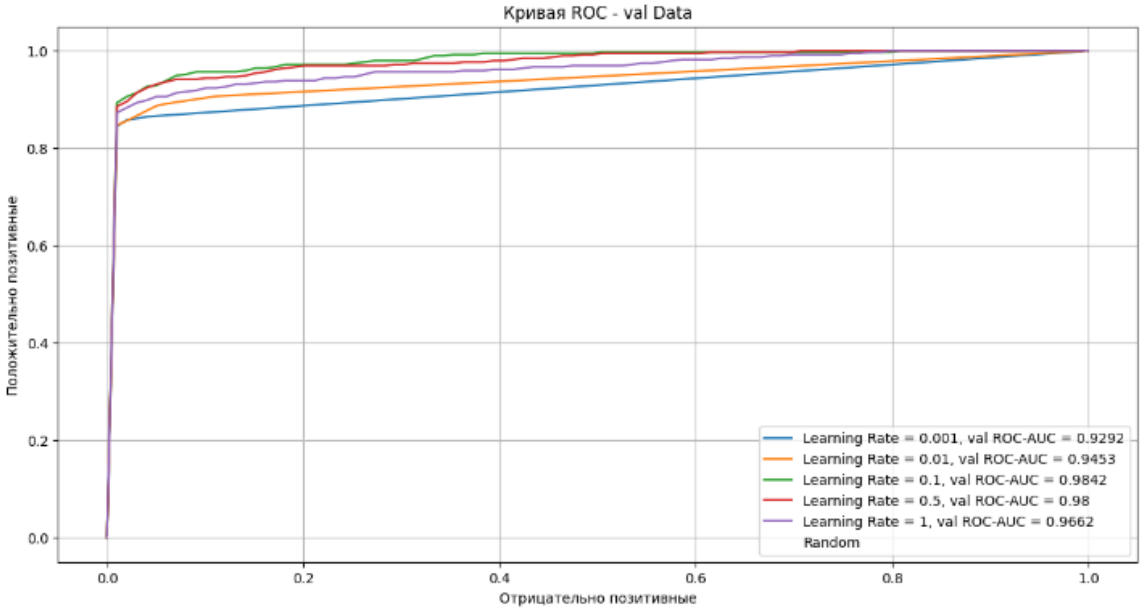

Рис. 5. Кривая ROC для градиентного спуска
Fig. 5. Gradient Descent ROC Curve

В первую очередь необходимо инициализировать модель с лучшими гиперпараметрами, полученными в ходе обучения модели. Далее предсказывается вероятность с помощью метода predict_proba. Затем рассчитывается значение ROC-AUC, предсказываются значения меток классов и получаются значения F1-Score, Precision, Recall.

В ходе проверки каждой модели были получены значения переменных и время, потраченное на проверку.

Логистическая регрессия

При гиперпараметре {“C”: 0,01, “penalty”: “l2”} получили следующие значения:

– значение ROC-AUC логистической регрессии на тестовом наборе: 0,975;
– значение F1-Score логистической регрессии на тестовом наборе: 0,598;
– значение Precision логистической регрессии на тестовом наборе: 0,479;
– значение Recall логистической регрессии на тестовом наборе: 0,796;
– затраченное время: 1,20 секунд.

Дерево решений

При гиперпараметре {“criterion”: “entropy”, “max_depth”: 3, “min_samples_leaf”: 1, “min_samples_split”: 2} получили следующие значения:

– значение ROC-AUC дерева решений на тестовом наборе: 0,931;
– значение F1-Score дерева решений на тестовом наборе: 0,82;
– значение Precision дерева решений на тестовом наборе: 0,803;
– значение Recall дерева решений на тестовом наборе: 0,837;
– затраченное время: 5,40 секунд.

Случайный лес

При гиперпараметре {“min_samples_split”: 7, “n_estimators”: 500} получили следующие значения:

– значение ROC-AUC случайного леса на тестовом наборе: 0,962;
– значение F1-Score случайного леса на тестовом наборе: 0,828;

– значение Precision случайного леса на тестовом наборе: 0,82;
– значение Recall случайного леса на тестовом наборе: 0,837;
– затраченное время: 1581,33 секунд.

Градиентный спуск

При гиперпараметре {“learning_rate”: 0,1, “max_depth”: 3, “subsample”: 0,7} получили следующие значения:

– значение ROC-AUC градиентного спуска на тестовом наборе: 0,969;
– значение F1-Score градиентного спуска на тестовом наборе: 0,824;
– значение Precision градиентного спуска на тестовом наборе: 0,812;
– значение Recall градиентного спуска на тестовом наборе: 0,837;
– затраченное время: 3,87 секунд.

На основе полученных результатов можно сделать вывод, что лучшей моделью машинного обучения для определения мошеннических транзакций стала модель, обученная с помощью алгоритма логистической регрессии, получившей значение ROC-AUC, равное 0,975.

Заключение

Была рассмотрена актуальная проблема обнаружения мошеннических транзакций в банковской сфере с использованием методов машинного обучения. Для решения данной задачи были использованы несбалансированные массивы данных. В ходе работы были проведены следующие этапы:

1. Предобработка данных (обезличивание, применение метода PCA, масштабирование с помощью RobustScaler и устранение асимметрии распределения).

2. Тестирование алгоритмов классификации. Для оценки моделей использовались метрики ROC-AUC, F1-мера. Среди четырех рассмотренных алгоритмов машинного обучения наибольшую эффективность показала логистическая регрессия (значение ROC-AUC составило 0,975 на тестовом массиве данных), в то время как остальные методы также имеют преимущества разного характера (например, случайный лес обеспечил высокую точность, но показал длительное время обработки). В некоторых случаях возможен компромисс между точностью и вычислительными ресурсами.

Таким образом, целесообразно применять машинное обучение для обнаружения и защиты банковских операций от мошенничества. Перспективы развития данной работы возможны за счет внедрения более глубоких нейросетевых архитектур, увеличения датасетов. Использование интеллектуальных технологий позволит автоматизировать процесс обнаружения мошенничества, снизить убытки и повысить уровень безопасности.

Список литературы

Аскаров Е.Ф., Хамитов Р.М. 2024. Использование временных рядов для прогнозирования мошеннических операций. *Экономика и предпринимательство*, 3(164): 1356–1359.

Григорьев А. 2023. Машинное обучение. Портфолио реальных проектов. Санкт-Петербург: Питер, 496 с.

Мартин Р. 2022. Чистая архитектура. Искусство разработки программного обеспечения. Санкт-Петербург: Питер, 352 с.

Марченко А.Л. 2023. Python: большая книга примеров. Издательство Московского университета, 361 с.

Окуньков С.В., Барулина М.А., Санбаев А.К. 2023. Мультиклассовая классификация на сильно несбалансированном наборе данных. Фундаментальная и прикладная медицина: материалы Международной конференции молодых ученых, Саратов, 105–106.

Орельен Ж. 2020. Прикладное машинное обучение с помощью Scikit-Learn, Keras и TensorFlow: концепции, инструменты и методы построения интеллектуальных систем, 2-е изд. Санкт-Петербург: ООО "Диалектика", 1520 с.

Плас Дж. В. Python для решения сложных задач: наука о данных и машинное обучение. Санкт-Петербург: Питер, 2021. 576 с.

Траск Э. 2025. Грокаем глубокое обучение. Санкт-Петербург: Питер, 352 с.

Хлобыстова А.О., Абрамов М.В. 2024. Публичность организации как уязвимость при проведении социоинженерной атаки. *Информационное общество*, 1: 85–93.

Chio K. 2020. Machine learning and security. Protecting systems with data and algorithms. Moscow: DMK Press, 388 p.

ICO Falcon Fraud Manager [Electronic resource]. URL: https://www.fico.com/en/products/fico-falcon-fraud-manager (date of request: 20.10.2025)

Ioffe L. 2024. Application of big data technology for fraud detection in financial transactions. *Universum: technical sciences*, 2(119): 6.

Kelleher J.D. 2019. Deep Learning. The Massachusetts Institute of Technology, 296 p.

Liu Yu., Li Ya., Xie D. 2024. Implications of imbalanced datasets for empirical ROC-AUC estimation in binary classification tasks. *Journal of Statistical Computation and Simulation*, 94(1): 183–203.

Madani A. 2023. Debugging Machine Learning Models with Python. Develop high-performance, low bias, and explainable machine learning and deep learning models. Birmingham: Packt Publishing Ltd, 344 p.

Omolara O., Agwubuo C., Onyeche S., Omotoyosi O., Kenneth N. and Olajumoke A. 2024. The impact of big data analytics on financial risk management. *International Journal of Science and Research Archive*, 12(02): 821–827.

Wang Y., Wang Q., Zhao L., Wang C. 2023. Differential privacy in deep learning: Privacy and beyond. *Future Generation Computer Systems*, 148: 408–424.

Ye J.X. 2023. A review of two-stage target detection algorithms based on deep learning. Internet Wkly, 18: 16–18.

Yuxi (Hayden) Liu. 2020. Python Machine Learning By Example. Third Edition. Build intelligent systems using Python, Tensor Flow 2, PyTorch, and scikit-learn. Birmingham: Packt Publishing Ltd, 526 p.

Zhu H., Zhou S.Y. 2023. A review of single-stage target detection algorithms based on deep learning. *Ind. Control. Comput*. 36: 101–103.

References

Askarov E.F., Xamitov R.M. 2024. Ispol`zovanie vremenny`x ryadov dlya prognozirovaniya moshennicheskix operacij [Using time series to predict fraudulent transactions]. *E`konomika i predprinimatel`stvo*, 3(164):1356–1359.

Grigor`ev A. 2023. Mashinnoe obuchenie. Portfolio real`ny`x proektov [Machine Learning. Portfolio of Real Projects]. Sankt-Peterburg: Piter, 496 s.

Martin R. 2022. Chistaya arxitektura. Iskusstvo razrabotki programmnogo obespecheniya [Clean Architecture. The Art of Software Development]. Sankt-Peterburg: Piter, 352 s.

Marchenko A.L. 2023. Python: bol`shaya kniga primerov [Python: a great book of examples]. Izdatel`stvo Moskovskogo universiteta, 361 s.

Okun`kov S.V., Barulina M.A., Sanbaev A.K. 2023. Mul`tiklassovaya klassifikaciya na sil`no nesbalansirovannom nabore danny`x [Multiclass classification on a highly imbalanced dataset]. Fundamental`naya i prikladnaya medicina: materialy` Mezhdunarodnoj konferencii molody`x ucheny`x, Saratov, 105–106.

Orel`en Zh. 2020. Prikladnoe mashinnoe obuchenie s pomoshh`yu Scikit-Learn, Keras i TensorFlow: koncepcii, instrumenty` i metody` postroeniya intellektual`ny`x sistem [Applied Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Methods for Building Intelligent Systems], 2-e izd. Sankt-Peterburg: OOO Dialektika, 1520 s.

Plas Dzh. V. Python dlya resheniya slozhny`x zadach: nauka o danny`x i mashinnoe obuchenie [Python for solving complex problems: data science and machine learning]. Sankt-Peterburg: Piter, 2021. 576 s.

Trask E`. 2025. Grokaem glubokoe obuchenie [Grok deep learning]. Sankt-Peterburg: Piter, 352 s.

Xloby`stova A.O., Abramov M.V. 2024. Publichnost` organizacii kak uyazvimost` pri provedenii socioinzhenernoj ataki [The organization's public nature as a vulnerability in a social engineering attack]. *Informacionnoe obshhestvo*, 1:85–93.

Chio K. 2020. Machine learning and security. Protecting systems with data and algorithms. Moscow: DMK Press, 388 p.

ICO Falcon Fraud Manager [Electronic resource]. URL: https://www.fico.com/en/products/fico-falcon-fraud-manager (date of request: 20.10.2025)

Ioffe L. 2024. Application of big data technology for fraud detection in financial transactions. *Universum: technical sciences*, 2(119): 6.

Kelleher J.D. 2019. Deep Learning. The Massachusetts Institute of Technology, 296 p.
Liu Yu., Li Ya., Xie D. 2024. Implications of imbalanced datasets for empirical ROC-AUC estimation in binary classification tasks. *Journal of Statistical Computation and Simulation*, 94(1): 183–203.
Madani A. 2023. Debugging Machine Learning Models with Python. Develop high-performance, low bias, and explainable machine learning and deep learning models. Birmingham: Packt Publishing Ltd, 344 p.
Omolara O., Agwubuo C., Onyeche S., Omotoyosi O., Kenneth N. and Olajumoke A. 2024. The impact of big data analytics on financial risk management. *International Journal of Science and Research Archive*, 12(02): 821–827.
Wang Y., Wang Q., Zhao L., Wang C. 2023. Differential privacy in deep learning: Privacy and beyond. *Future Generation Computer Systems*, 148: 408–424.
Ye J.X. 2023. A review of two-stage target detection algorithms based on deep learning. Internet Wkly, 18: 16–18.
Yuxi (Hayden) Liu. 2020. Python Machine Learning By Example. Third Edition. Build intelligent systems using Python, Tensor Flow 2, PyTorch, and scikit-learn. Birmingham: Packt Publishing Ltd, 526 p.
Zhu H., Zhou S.Y. 2023. A review of single-stage target detection algorithms based on deep learning. *Ind. Control. Comput.* 36: 101–103.

Конфликт интересов: о потенциальном конфликте интересов не сообщалось.
Conflict of interest: no potential conflict of interest related to this article was reported.

Поступила в редакцию 24.10.2025
Поступила после рецензирования 13.01.2026
Принята к публикации 20.01.2026

Received October 24, 2025
Revised January 13, 2026
Accepted January 20, 2026

ИНФОРМАЦИЯ ОБ АВТОРАХ

Хамитов Ренат Минзашарифович, кандидат технических наук, доцент, доцент кафедры информационных технологий и интеллектуальных систем, Казанский государственный энергетический университет, г. Казань, Россия

Куценко Светлана Мунавировна, кандидат педагогических наук, доцент, доцент кафедры информационных технологий и интеллектуальных систем, Казанский государственный энергетический университет, г. Казань, Россия

Салтанаева Елена Андреевна, кандидат технических наук, доцент кафедры информационных технологий и интеллектуальных систем, Казанский государственный энергетический университет, г. Казань, Россия

INFORMATION ABOUT THE AUTHORS

Renat M. Khamitov, Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Information Technologies and Intelligent Systems, Kazan State Power Engineering University, Kazan, Russia

Svetlana M. Kutsenko, Candidate of Pedagogical Sciences, Associate Professor, Associate Professor of the Department of Information Technologies and Intelligent Systems, Kazan State Power Engineering University, Kazan, Russia

Elena A. Saltanaeva, Candidate of Technical Sciences, Associate Professor of the Department of Information Technologies and Intelligent Systems, Kazan State Power Engineering University, Kazan, Russia